

## MUSIC RETRIEVING DEVICE AND METHOD, AND STORAGE MEDIUM RECORDED WITH MUSIC RETRIEVING PROGRAM

Publication number: JP2000347659 (A)

Also published as:

Publication date: 2000-12-15

□ JP3631660 (B2)

Inventor(s): NISHIHARA YUICHI; KOSUGI NAOKO; SAKATA TETSUO;  
YAMAMURO MASASHI; UMEDA MASAYOSHI; KONYA SEIICHI  
+

Applicant(s): NIPPON TELEGRAPH & TELEPHONE +

Classification:

- International: G06F17/30; G10H1/00; G10H1/40; (IPC1-7): G06F17/30;  
G10H1/00; G10H1/40

- European:

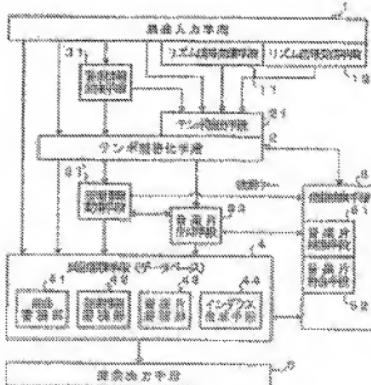
Application number: JP20000008458 20000118

Priority number(s): JP20000008458 20000118; JP19990082722 19990326

### Abstract of JP 2000347659 (A)

PROBLEM TO BE SOLVED: To precisely and quickly retrieve music intended by a retriever out of a database, by inputting freely a portion or the whole of the music without making this retriever be aware of a phrase, a melody, a tempo and the like.

SOLUTION: A tempo of input music or musical information converted from the input music is normalized by a tempo normalizing means 2 to be stored in a database by a music accumulating means 4, in database construction. The tempo of the music or the musical information converted from the music, which is a retrieving key, is normalized by the tempo normalizing means 2, similarly retrieval is carried out by a music retrieving means 5 based thereon, and the music as a retrieved result is output by a music outputting means 6, when retrieved.



(19) 日本国特許庁 (JP)

## (12) 公開特許公報 (A)

(11) 特許出願公開番号  
特開2000-347659  
(P2000-347659A)

(13) 公開日 平成12年12月15日 (2000.12.15)

(51) Int CL'      細別記号  
G 10 H 1/00      1 0 2  
G 06 F 17/30  
G 10 H 1/40

G 10 H

1/00

1/40

15/40

15/403

P 1

G 10 H

1/10

3 7 0 E

3 5 0 C

テ-33-1 (参考)

1 0 2 Z

3 7 0 E

3 5 0 C

審査結果 未請求 請求項の数14 O.L (全 12 頁)

(21) 出願番号 特願2000-8458 (P2000-8458)  
(22) 出願日 平成12年1月18日 (2000.1.18)  
(31) 優先権主張番号 特願平11-82722  
(32) 優先日 平成11年3月26日 (1999.3.26)  
(33) 優先権主張国 日本 (JP)

(71) 出願人 000004226  
日本電信電話株式会社  
東京都千代田区大手町二丁目3番1号  
(72) 発明者 西原 祐一  
東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内  
(72) 発明者 小杉 尚子  
東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内  
(74) 代理人 100087848  
弁理士 小笠原 審義 (外1名)

最終頁に続く

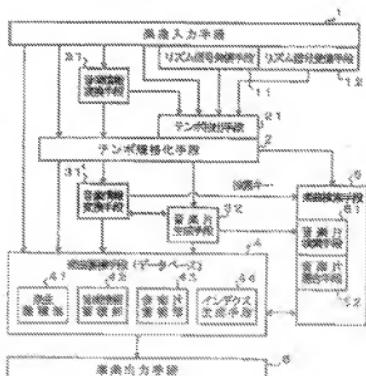
(54) 【発明の名称】 音楽検索装置、音楽検索方法および音楽検索プログラムを記録した記録媒体

## (57) 【要約】

【課題】 検索者が、フレーズ、調、テンポ等を意識することなく、自由に楽曲の一音または全体を入力することにより、検索者が登録した楽曲をデータベース中から精度よく、かつ高速に検索できるようにする。

【解決手段】 データベースの構成では、入力された楽曲もしくは楽曲から変換された楽曲情報のテンポを、テンポ獲得化手段2により算出後、楽曲登録手段4によりデータベースに格納する。検索の部には、テンポ獲得化手段2により検索キーとなる楽曲もしくは楽曲から変換された音楽情報のテンポを算出化し、それをもとに検索登録手段5により翻訳検索を行い、楽曲出力手段6により検索結果の楽曲を出力する。

## 本発明の構成概要



## 【特許請求の範囲】

【請求項1】 楽曲を検索する検索装置において、楽曲を入力する楽曲入力手段と、上記楽曲入力手段により入力された楽曲のテンポを規格化するテンポ規格化手段と、少なくとも上記テンポ規格化手段によりテンポ規格化された楽曲を蓄積する楽曲蓄積手段と、検索キーとして入力された楽曲の一部または全体から上記テンポ規格化手段によりテンポ規格化された楽曲の一部または全体をもとに、上記楽曲蓄積手段により蓄積された楽曲の一部または全体の中から類似した楽曲を検索する楽曲検索手段と、上記楽曲検索手段により検索された楽曲を出力する楽曲出力手段とを有することを特徴とする音楽検索装置。

【請求項2】 請求項1記載の音楽検索装置において、上記楽曲入力手段により入力された楽曲を音楽情報に変換する音楽情報変換手段を有し、上記テンポ規格化手段は、楽曲を上記音楽情報変換手段により音楽情報に変換した後でテンポ規格化を行うことを特徴とする音楽検索装置。

【請求項3】 請求項1または請求項2記載の音楽検索装置において、上記楽曲入力手段は、一定の時間間隔もしくは決められた周期に従った時間間隔で信号を発信するリズム信号発信手段、または、楽曲入力者が入力中の楽曲のリズムを合わせて入力できるようなリズム信号受信手段を有することを特徴とする音楽検索装置。

【請求項4】 請求項1ないし請求項3のいずれかに記載の音楽検索装置において、上記テンポ規格化手段は、上記楽曲入力手段により入力された楽曲、上記音楽情報に変換手段により変換された楽曲蓄積手段、上記リズム信号発信手段により発信された信号、上記リズム信号受信手段により受信された信号の少なくとも一つを用いてテンポを抽出するテンポ抽出手段を有し、抽出されたテンポを用いてテンポ規格化を行うことを特徴とする音楽検索装置。

【請求項5】 請求項1ないし請求項4のいずれかに記載の音楽検索装置において、上記楽曲検索手段は、テンポ規格化された検索キーの一一部または全体をもとに、上記楽曲蓄積手段により蓄積された楽曲の一部または全体を検索する際に、テンポ規格化された楽曲の検索キーの一一部または全体と、テンポ規格化された楽曲の検索キーの一一部または全体と回し検索の、検索対象となるテンポ規格化された楽曲の一部または全体との類似度を調べることを特徴とする音楽検索装置。

【請求項6】 信号処理手段と、し請求項5のいずれかに記載の音楽検索装置において、上記テンポ規格化された楽曲もしくはテンポ規格化された音楽情報を用いて、音楽信号の分割を行う算出手段を有することを特徴とする音楽検索装置。

【請求項7】 請求項6に記載の音楽検索装置において、上記算出手段は、音楽片を生成する際にすべての

音楽片の長さを一定にすることを特徴とする音楽検索装置。

【請求項8】 請求項6または請求項7記載の音楽検索装置において、上記音楽片生成手段は、音楽片を生成する際に、同じ楽曲に対する音楽片は、一定の時間間隔で音楽片が開始するように生成する、あるいは音楽片が音符の始まりから開始する、または音の高さの変動もしくは変動率がある一定の閾値を超えた点から開始するよう生成することを特徴とする音楽検索装置。

【請求項9】 請求項1ないし請求項8のいずれかに記載の音楽検索装置において、上記楽曲蓄積手段により蓄積する前に音楽情報への変換を行う音楽情報変換手段を有することを特徴とする音楽検索装置。

【請求項10】 請求項1ないし請求項9のいずれかに記載の音楽検索装置において、楽曲を蓄積する際に、検索を高速化するためのインデックスを生成するインデックス生成手段を有し、上記楽曲検索手段は、検索を行う際にインデックスを用いることを特徴とする音楽検索装置。

【請求項11】 請求項1ないし請求項10のいずれかに記載の音楽検索装置において、上記楽曲検索手段は、検索を行う際に、音楽片同士の検索を行い、その結果の音楽片同士の関係から最終的な検索結果を導く音楽片検索手段を有することを特徴とする音楽検索装置。

【請求項12】 請求項11記載の音楽検索装置において、上記楽曲検索手段は、検索キーの検索対象同士の関係と、検索結果の音楽片同士の関係の整合性を調べた後、最終的な検索結果を導く音楽片整合手段を有することを特徴とする音楽検索装置。

【請求項13】 楽曲を検索する検索方法において、楽曲を入力する過程と、入力された楽曲のテンポを規格化する過程と、少なくとも、テンポ規格化された楽曲を蓄積する過程と、検索キーとして入力された楽曲の一部または全体をもとに、テンポ規格化された楽曲の一部または全体から得られた、テンポ規格化された楽曲の一部または全体と回し検索の、検索対象となるテンポ規格化された楽曲の一部または全体との類似度を調べることを特徴とする音楽検索方法。

【請求項14】 検索機によって楽曲を検索するために用いるプログラムを記録したプログラム記録媒体であって、楽曲を入力する処理と、入力された楽曲のテンポを規格化する処理と、検索キーとして入力された楽曲の一部または全体をもとに、あらかじめテンポ規格化されて蓄積された検索対象となる楽曲の一部または全体とから、回し検索した楽曲を検索する処理と、検索された楽曲を出力する処理とを、記録機に実行させるプログラムを記録した記録媒体。

【別掲の詳細な説明】  
【図1】

【発明の属する技術分野】本発明は、音楽検索に係わり、特に、楽曲入力手段によって入力もしくは指定された楽曲の一部もしくは全体を検索キーとして、データベースに蓄積された楽曲の中から、検索キーと類似する楽曲の一部もしくは全体を検索する装置およびその方法に関する。

### 【0002】

【従来の技術】音楽の類似検索を行う場合、従来方法においては音符がマッチングの処理単位として用いられていた。この場合、音符とは、音の高さの変動もしくは変動率がある一定の範囲内に収まっている音の連続のこととを言う。また、各音符は、音符情報として音の高さ(音名)、音の長さ(音長)、時間的に前後する音符との音名の差(相対音高)、時間的に前後する音符との音長の比(相対音長)などを保有している。

【0003】従来は、従来方法による音楽検索のうち、音符を単位としたDIPマッチングを使用するものの原理を説明する例である。まず、データベース構築のフェーズでは、検索対象とする楽曲を入力し(ステップS101)、音符情報を変換する(ステップS102)。それをデータベースへ格納する(ステップS103)。データベースへ格納した楽曲の検索は、検索キーとなる楽曲を入力し(ステップS111)、それを音符情報を変換する(ステップS112)。次に、データベース中の全楽曲の音符情報を検索キーから得られた音符情報をついて、データベース中の曲数分、D.P(Dynamic Programming)マッチングを実施し(ステップS113)、もともと類似した楽曲の候補を検索結果として出力する(ステップS114)。

【0004】以上のように、音楽の類似検索を行う場合、従来方法においてはDIPマッチングが使われていた(参考文献1、2)。

【参考文献1】樋田勝也、後藤真理、利岡洋一:「WWW上の歌による曲検索システム」、電子情報通信学会技術報告書、SP-97-103、1998。

【参考文献2】鶴山哲也、高橋幸典:「ハミング歌謡を手術りとするメロディ検索」、電子情報通信学会論文誌、Vol.177-4/1998 pp.1543-1551、1994。

DIPマッチングにおいては、検索キーと、データベース中の楽曲の一部または全体(検索結果曲)の対応関係が求められ、対応関係にある検索キーの音符と、候補部分曲の音符との間の、音符間の距離によって評価が計算される。この検索キーとの間の相似度が最大になるような距離を部分曲を、距離の順序によって探し出すのがDIPマッチングである。

### 【0005】

【検索を実現しようとする課題】従来方法の音符をデータベースの単位とした検索方法では、一般に、許容できる音符誤りの数を斟酌している。

【0006】また、検索方法のDIPマッチングでは、計

算量を抑制するために、検索キーの音符列と、類似度計算対象となるデータベース中の楽曲の音符列との間の音符数の差異に制限を設ける。

【0007】したがって、音楽検索をする場合に、検索キー入力の際に、検索目標とするデータベース中の楽曲に對して一定数以上に音符の插入・削除を行ってしまうと検索精度が悪くなるという問題があつた。

【0008】とえ検索キー入力の際に、音符の挿入・削除が行われなかつたとしても、検索キーから音符情報を抽出する際に、音符の挿入・削除が生じる可能性も大きく、そのため検索精度が悪いという問題があつた。【0009】また、D.Pマッチングにおいては、インデックスを利用するにとめできず、データベース中の全曲に對して全検索を施す必要があるために、たとえ計算量を抑制するために許容できる音符誤りの量を制限したとしても、やはり検索に時間がかかるという問題があつた。

【0010】一方、音楽片を使用した検索の場合、インデックスを利用するために検索時間の大半を削減することが可能であった。

【0011】図10は、音楽片を使用する音楽検索の原理を説明する例である。この手法は、特願平10-341754号(音楽検索装置、音楽検索方法および音楽検索プログラムを記録した記録媒体)で、本発明者が権利を取っているものである。

【0012】まず、データベースの構築では、検索対象とする楽曲を入力し(ステップS201)、入力した楽曲を音符情報を変換し(ステップS202)、その音符情報を複数の音楽片に分割し(ステップS203)。音楽片が可選のように、インデックスを生成し(ステップS204)、音楽片の情報をインデックスとともにデータベースに格納する(ステップS205)。

【0013】次に、楽曲の検索では、検索キーとなる楽曲を入力し(ステップS211)、入力した楽曲を音符情報を変換し(ステップS212)、音符情報を複数の音楽片に分割する(ステップS213)。音楽片の一つ一つを検索キーとして、インデックスを用いてデータベースに蓄積された音楽片から類似度をもを検索する(ステップS214)。検索結果として検出された音楽片同士の関係などを算出する(ステップS215)。検索キーとなった音楽片同士の関係などを調べることで、検出された音楽片を含むデータベース中の楽曲の一部または全曲との間の類似度を算出する(ステップS216)。検索キーとなった音楽片の一部または全曲と類似度の高いデータベース中の楽曲の一部または全曲を検索結果として出力する(ステップS217)。

【0014】音楽片検索を行なう場合には、類似度を計算する音楽片同士の音符の数が一致しているために、DIPマッチングを利用すると検索結果が得られなくなる。インデックスを利用した検索が検索が可能である。

【0015】しかし、音楽片検索をする際は、音符情報が認められないために、音楽片の長さ分だけ、音符の挿入・削除のない区間が必要であるという制限があった。すなわち、音楽片の長さ分だけ、音符の挿入・削除のない区間がなければ、検索精度が悪いという問題点があった。

【0016】また、音符情報を必ずしも必要としない方法としては、音楽情報への変換を行った後、類似検索を行いういう手法がある。

【0017】図11は、音符情報を使用しない音楽検索の原理を説明する図である。この手法は、特願平10-329131号(音楽情報検索装置、音楽情報検索装置、音楽情報検索方法、音楽情報検索方法およびそれらのプログラムを記録した記録媒体)で、本発明等が提案しているものである。

【0018】まず、データベースの構築において、楽曲を入力し(ステップS301)，入力した楽曲から音楽的な特徴(特徴量)を抽出することが可能な音楽情報に変換し(ステップS302)，それをデータベースへ格納する(ステップS303)。

【0019】楽曲の検索において、検索キーとなる楽曲を入力し(ステップS311)，入力した楽曲を同等な音楽情報への変換を行い(ステップS312)。変換した音楽情報をもとに、データベース中の楽曲について類似度を実現する(ステップS313)。ここで、類似検索として、検索キーの楽曲の特徴量をデータベースに格納された楽曲の特徴量と比較し検索し、検索は距離度数を用いて距離の近いものと距離度の高いものとする類似度計算を行う。類似したものを含む楽曲の候補を検索結果として出力する(ステップS314)。

【0020】しかし、このような音楽情報を利用する方法においては、検索キーと、データベース中の曲とのテンポが異なる場合に、検索精度が悪いという問題点があつた。

【0021】本発明は、上記の点に鑑みなされたもので、データベース中の楽曲および検索キーのテンポを規格化することにより、音符を削除するのではなく、規格化された時間を単位として検索を可能とすることによって音符数によりて影響を受けない楽曲の検索装置および方法を提供することを目的とする。本発明においては、DMPマッチングを使用するがいため、インデックスを用いた検索を行うことが可能であり、検索の検索が可能である。また、検索を行う前にアンボを規格化するためには、検索キーとデータベース中の曲とのテンポが違う場合に、長い検索精度を得ることが可能である。

【0022】

【楽曲を検索するための手段】図1および図2は、本発明による音楽の類似検索の装置を説明するための図。図3は、本発明の構成図である。

【0023】本発明は、音楽検索において、データベースに入力された楽曲もしくは楽曲から変換された音楽情報のテンポを規格化した上でデータベースに蓄積し、また、検索キーとして入力された楽曲もしくは楽曲から変換された音楽情報のテンポを規格化した上で、音楽の類似検索を行うことを主な特徴とする。

【0024】まず、音楽情報への変換を行わない検索におけるデータベースの構築では、楽曲入力手段1を用いて楽曲が入力される(ステップS10)。入力された楽曲は、テンポ規格化手段2によってテンポが規格化される(ステップS11)。

【0025】また、音楽情報への変換を行なう検索の場合のデータベースの構築では、楽曲入力手段1を用いて楽曲が入力された後(ステップS30)、音楽情報変換手段31により音楽情報に変換され(ステップS31)。その後、テンポ規格化手段2によってテンポが規格化される(ステップS32)。

【0026】テンポ規格化された楽曲あるいはテンポが規格化された音楽情報は、楽曲蓄積手段4により、音楽情報蓄積手段42にデータベースとして蓄積される(ステップS12, 533)。

【0027】テンポ規格化された楽曲あるいはテンポが規格化された音楽情報は、楽曲蓄積手段4により蓄積される前に、音楽情報変換手段31による音楽情報への変換、音楽片生成手段32による音楽片41の分割の過程を経てから、蓄積されることもある。さらに、蓄積された楽曲、音楽情報、音楽片に対して、インデックス生成手段44により、検索を高速化するためのインデックスが作成されることもある。

【0028】楽曲蓄積手段4には、テンポ規格化される前の楽曲を楽曲蓄積手段41に蓄積することも可能であり、検索の結果として、テンポ規格化される前の楽曲を出力するにも可能である。

【0029】楽曲の検索においては、まず、楽曲入力手段1を用いて楽曲の一部または全体が入力もしくは検定される(ステップS20, S40)。入力された楽曲は、テンポ規格化手段2によりテンポが規格化されるか(ステップS21)、音楽情報変換手段31により音楽情報が変換された後(ステップS41)、テンポが規格化される(ステップS42)。テンポ規格化された検索キーは、さらに、音楽情報変換手段31により音楽情報へ変換され、音楽片生成手段32により音楽片41が分割されることもある。

【0030】上記処理を繰り返すキーは、楽曲蓄積手段4へ送られる。音楽検索手段42では、検索キーへ対応する楽曲の一時もしくは全曲を、データベースに蓄積された楽曲が並列処理する(ステップ530, 540)。データベースへインデックスが生成されない場合は、データベースを用いて検索を行う。

【0031】音楽片を用いた検索では、音楽片検索手段

5.1により、検索キーとして送られてきた音楽片の少なくとも一つを用いて、データベースに蓄積された音楽片から類似するものを検索し、楽曲出力手段1により、検索結果の楽曲を出力する(ステップS2.3、S4.4)。検索キーとして送られてきた音楽片の複数を用いて検索する場合には、検索キーのそれぞれの音楽片に類似するデータベース中の音楽片を検索した後、音楽片整合手段5.2により、音楽片同士の整合性を調べ、最終的な検索結果とすることもある。

【0032】なお、音楽片の整合については、上記の関連する特許出願「特願平10-341754号」の明細書等に記載されている手法を採用する。具体的には、検索結果として検出された音楽片同士の関係と、検出された音楽片のそれぞれの検索キーとなった音楽片同士の関係を調べ、また、検出された音楽片の前後の音符と、検出された音楽片の検索キーとなった音楽片の前後の音符の関係を調べることで、検索キーとなった楽曲の一部または全部と、検出された音楽片を含むデータベース中の楽曲の一部または全部との間の類似度を調べる。

【0033】楽曲入力手段1は、一定の時間間隔もしくは決められた順序に従った時間間隔で信号を発するリズム信号発信手段11を有することもある。特に、検索キーとして入力される楽曲に則しては、テンポ規格化が十分な精度で行われない可能性がある。そこで、リズム信号発信手段11によって一定の時間間隔もしくは決められた順序に従った時間間隔で信号を発信させ、その信号に合わせて楽曲入力者が楽曲の入力をを行い、入力された楽曲と、リズム信号発信手段11により発信される信号とを同時にテンポ規格化手段2に入力させることによって、正確なテンポ規格化を行うことを可能とする。

【0034】あるいは、楽曲入力手段1は、リズム信号発信手段12を有することもある。これは、楽曲入力者が楽曲の演奏(映像)に合わせて発するリズム信号を検出可能とする手段で、入力された楽曲と、リズム信号発信手段12により受信される信号を同時にテンポ規格化手段2に入力させることによって、正確なテンポ規格化を行うことを可能とする。

【0035】テンポ規格化手段2は、楽曲入力手段1により入力された楽曲、音楽情報変換手段3により変換された音楽情報、もしくは、リズム信号発信手段11により受信された信号から、テンポが抽出手段21を有するにともなる。この場合、テンポ規格化手段2は、テンポが抽出手段21によって抽出されたテンポを用いて、テンポ規格化を行う。

【0036】以上に、本発明の作用を説明する。図4は、テンポ規格化について説明するための図である。楽曲に就いては、テンボによって、音楽片は「拍子」の時間的な長さが測定される。図4(a)の「その楽曲」に示されているように、楽曲はそれぞれテンボを持っており、このテンボは、楽曲開頭の間に一致するところもある。また、同じ楽曲であっても、一つ一つの楽章によって異なるところもある。また、同じ楽曲であっても、曲の部分によって異なるところもある。

【0037】テンボを規格化するということは、図4(b)に示すように、すべての楽曲のすべての訊みの、すべての部分において、拍子の長さが一定の時間的な長さになるように、楽曲に変換を施すことである。

【0038】図5は、マッチングの単位を規格化するための図である。テンポ規格化を行わない場合、主に音符がマッチングの単位とされてきた。しかし、図5(A)に示すように、音符の挿入・削除が生じると、比較する音符同士の対応関係がずれるという問題があった。DPマッチングは、音符の挿入・削除があった場合の比較する音符同士の対応関係のズれを修正するための方法である。

【0039】実際にDPマッチングを用いて検索を行う場合、許容する音符の挿入・削除の量には制限を設ける必要があるため、音符の挿入・削除が検索精度に大きな影響を与える。また、DPマッチングを用いた検索では、検索を高速に行うためのインデックスの作成などもできないため、多くの検索時間を必要とする。

【0040】一方、テンボ規格化を行った場合、図5(B)のように時間を単位としたマッチングが可能となる。この場合、音符の挿入・削除があったとしても、対応関係にずれが生じることがなく、そのことにより検索精度が影響されることはない。したがって、精度の高い検索が可能となる。また、テンボ規格化された楽曲を、音楽片に分割し、インデックスを作成することにより、検索を高速に行うこと也可能である。

【0041】本発明に関連する技術として、特願平10-329131号、特願平10-341754号に開示する技術がある。本発明は、データベース構造の際に、データベースに入力する楽曲のテンボ規格化を行い、検索キーとなる楽曲の入りの際もテンボ規格化を行なう。が、先の発明と主に異なる。特に、本発明は、特願平10-341754号に開示する発明とは、特に音符単位での検索を行わず、時間軸を軸とした検索を行なうに異なる。

【0042】【説明の実施の形態】【第1の実施の形態】本発明の第1の実施の形態について説明する。データベースに入力するための楽曲として、MIDIデータを使用する。現在ではカラオケ用のデータを始め、多くの楽曲がMIDIデータとして販売されている。

【0043】MIDIデータでは、音符情報が拍子を基準として記述されている。そして、拍子の時間長さの長さが、テンポ情報として、音符情報と共に記述されている。したがって、規格化された楽曲として、楽曲を標準とした時間を用いれば、データベース中の楽曲のナ

ンボ規格化が実現される。

【0044】データベースに格納するにあたり、元のMIDIデータは、まず、音楽片に分割される。分割の方法としては、音楽片の長さが、規格化された時間軸上ですべて同じになるようにし、図7(A)に示すように、すべての音符の始まりを先頭とする音楽片が生成されるようとする方法がある。あるいは、図7(B)に示すように、音楽片が一定の時間間隔で始まるように分割する方法もある。

【0045】また、音楽片へ分割した後、音楽情報(特徴量)として、音高推移ベクトルを生成する。これは、図7に示すように、規格化された時間軸を等間隔の目盛で刻み、各自の位置で演奏されている音の高さを並べてできるベクトルである。音符級での処理上、音の高さは、例えばMIDIで使われている音名を表す数値や、頭波数などに置き換える。

【0046】音楽情報(特徴量)としては、上記のほかに、例えば、以下のものを生成することができる。

【0047】① 音高推移ベクトルA<sub>1</sub> := (a<sub>11</sub>, a<sub>12</sub>, ..., a<sub>1n</sub>)としたとき、ベクトルの要素の平均値 m<sub>1</sub> := (a<sub>11</sub> + a<sub>12</sub> + ... + a<sub>1n</sub>) / n を計算し、各要素からこの平均値を引いた値を要素とするベクトルA<sub>2</sub> := (a<sub>11</sub> - m<sub>1</sub>, a<sub>12</sub> - m<sub>1</sub>, ..., a<sub>1n</sub> - m<sub>1</sub>) を特徴量とする。検索キーの楽曲と、データベース中の楽曲の調が異なっていても、このベクトルは類似することが期待される。

【0048】② 音高推移ベクトルA<sub>3</sub> := (a<sub>11</sub>, a<sub>12</sub>, ..., a<sub>1n</sub>) に対し、最初の要素a<sub>11</sub>を基準にし、各要素からこの基準値を引いた値を要素とするベクトルA<sub>4</sub> := (0, a<sub>12</sub> - a<sub>11</sub>, ..., a<sub>1n</sub> - a<sub>11</sub>) を特徴量とする。検索キーの楽曲と、データベース中の楽曲の調が異なっていても、このベクトルは類似することが期待される。

【0049】③ 音高推移ベクトルA<sub>5</sub> := (a<sub>11</sub>, a<sub>12</sub>, ..., a<sub>1n</sub>) に対し、各要素から前の要素の値を引いた値を要素とするベクトルA<sub>6</sub> := (a<sub>11</sub> - a<sub>12</sub>, a<sub>12</sub> - a<sub>13</sub>, ..., a<sub>1n</sub> - a<sub>1n-1</sub>) を特徴量とする。検索キーの楽曲の調が少しずつ進歩してしまう場合でも、データベース中の楽曲と、このベクトルは類似することが期待される。

【0050】これらの音楽情報、データベースに格納された後、楽曲を高確率化するためのインデックスを生成する。インデックスの生成方法としては、K-treesなどを使う。なお、K-treesに従事する参考文献としては以下のものがある。

「参考文献5」より引用。"DEEPS: A Database Index Structure for Spoken Language Processing", 田中和也著、日本語学会編『国際言語学会論文集』、日本言語学会、pp.47-57、1984。  
検索キーとなる楽曲の入力においては、ハッシュによる

入力、楽曲の演奏による入力、MIDI+キー+ボードの演奏による入力などの方法がある。

【0051】楽曲の入力の際には、リズム信号発信装置として、電子的なメトロノームを用いる。この電子的なメトロノームは、指定したテンポで音や光、コンピュータディスプレイ上の文字、图形などを鳴らすことができるようになっており、楽曲入力者が、その発する信号に合わせて楽曲の入力ができるようになっている。この電子的なメトロノームは、一定の時間間隔で信号を発するだけでなく、例えば、小節の先頭の拍は音量や、音の色、光の強さ、ディスプレイ上の文字の種類、图形の種類を、他の拍とは異なるように発したり、また例えば「サンバのリズム」など、入力する曲の種類に合わせたりズム負けを発することができる。また、この信号を発するテンポであるが、楽曲入力者が楽曲の入力をしやすいように、遅くしたり速くしたりを自由に設定することができる。

【0052】あるいは、楽曲入力の際には、リズム信号受信装置を使うこともある。これは、手拍子、足踏み、手の上下の動き、体の擺れ、マウスのクリック、キーボードの押下など、楽曲入力者が楽曲入力の際に意識的/無意識的に発する動き、信号などを検知できるようにした装置である。

【0053】入力された楽曲と、上記リズム信号受信装置から得られる信号を使って、検索キーとなる楽曲のテンポ規格化を行う。

【0054】音楽片への分割、音楽情報の変換については、MIDI+キー+ボードから入力される楽曲については、データベースへの楽曲の入力の場合と同じである。

【0055】ハミングによる入力、楽曲の演奏による入力では、楽曲入力手段1によりPCM形式のデータとして楽曲が入力される。そして、このデータから音高情報が抽出される。音高情報の抽出に関するPCTM形式のデータから音符情報を抽出する從前方法に準じる。

【参考文献4】新藤高水、今井正和、井原伸士:「楽曲の自動認識」、計測自動制御学会論文誌、Vol.20 No.10 pp.340-345、1984。【参考文献5】水野正典、藤本正義、高島勝彦、鶴治七郎:「パーソナルコンピュータによる楽曲の自動認識」、情報処理学会第35回全国大会、55-59、1997。

すなわち、まず、FFT等、自己相關処理を用いて周波数分析を行い、ピーア抽出、各ピーアのパワーメ定などを行ふ。さらに、ノイズの除去、ピッチャーフィットの検定、高調波成分の基準成分への吸収などをを行い、時間軸上の利害に對して、どの時点での音が被覆されているかを検出する。音楽情報を抽出された後、音楽片への分割、音楽情報を生成の仕方について、データベースへの楽曲の入力の場合と同じである。

【0056】なお、音楽情報を生成の過程を二段階に分け、連れてデータを規格化を行う場合もある。例えば、ま

す音品情報を抽出した後、テンポを規格化し、その後、音高ベクトルなど特徴量の生成を行なう場合もある。

【0057】検索については、まず、検索キーとして人力された楽曲から生成された音楽片のそれぞれに類似する音楽片を、データベース中から検索する（音楽片検索）。その際には、検索キーの音楽片とデータベース中の音楽片との距離を、音高ベクトルなどそれぞれの特徴量に対して求める。なお、この距離は、ユークリッド距離やマンハッタン距離など適当な距離を特徴量ごとに設定して求める。また、検索を高効率化するため、インデックスを用いて検索を行う。

【0058】音楽片同士の距離（非類似度）は、各特徴量の距離に適当な重みを加味して何らかの方法で求めたものとして表す。詳しくは、以下のとおりである。用いている特徴量の集合を  $F = \{F_1, F_2, \dots, F_n\}$  とする。音楽片  $K$  と音楽片  $X$  との特徴量  $F$  についての距離を  $d(F(K), X)$ 、重みを  $W$  とすると、音楽片  $K$  と音楽片  $X$  との総合的な距離  $D(K, X)$  を何らかの方法で求めるのは、次のような関数  $F$  を用いて行うことになる。

【0059】

$$D(K, X) = F([W]), \quad d(F(K, X))$$

（ただし、 $F \in F\text{SPACE}$ ）

この関数  $F$  の具体的な例として、幾何形の関数を用いるとする。

$$D(X, K) = \text{sum}(W \times d(F(K, X)))$$

となる（ただし、 $F \in F\text{SPACE}$  であり、 $\text{sum}$  は  $F \in F\text{SPACE}$  についての総和をとる）。

【0060】関数  $F$  として幾何形ではなく、例えば次のような最小値をとる関数を用いることもできる。

【0061】

$$D(X, K) = \min(W \times d(F(K, X)))$$

（ただし、 $F \in F\text{SPACE}$ 、 $\min$  は  $F \in F\text{SPACE}$  についての最小値をとる）

また、上記幾何形と最小値との混合を用いることもできる。すなはち、特徴量  $F \in F\text{PACE}$  を 2 種類 ( $F\text{PACE}_1, F\text{PACE}_2$ ) に分けて、 $F\text{PACE} = F\text{PACE}_1 \sqcup F\text{PACE}_2$  であるとする。

$$D(X, K) = \text{sum}(W_1 \times d(F_1(K, X)), W_2 \times d(F_2(K, X)))$$

のような関数を用いて複合距離を求めるところができる。ここで、 $W_1 \in F\text{PACE}_1, W_2 \in F\text{PACE}_2$  である。

【0062】ここでは、音楽片検索の距離（非類似度）の求め方として、距離和を用いる場合、最小値を用いる場合、幾何形と最小値の混合を用いる場合の例を説明したが、もちろんこの他にも距離和の求め方はいろいろあるが、データベース中の音楽片のうち、検索キーの音楽片との距離が近いもの、すなはち類似度の高いものが音楽片検索の結果となる。

【0063】検索キーから音楽片が複数生成される場合、検索キーの音楽片のそれぞれに対する音楽片検索の結果に対して音楽片検索の整合性を調べ、複数の音楽片を組み合わせた場合の類似度を計算し、最終的な検索結果とすることができる。

【0064】今、検索キーの音楽片  $K_1, \dots, K_n$  ( $p \leq n$ ) に対し、データベース中の音楽片  $X, Y$  がそれぞれ類似していたとする ( $K_i \sim X, K_i \sim Y$ )。音楽片  $X, Y$  が検索キーの音楽片  $K_1, \dots, K_n$  に対して整合性を持つということは、以下の条件を満たすことである。

【0065】1) 音楽片  $X, Y$  は同じ曲に属さなければならない。なお、以降ではこの候補曲を  $A$  とし、音楽片  $X, Y$  をそれぞれ  $A_1, A_2$  とおく。

【0066】2) 音楽片  $A_1, A_2$  の時刻的順序は、対応する検索キーの音楽片 ( $K_1, \dots, K_n$ ) の時刻的順序と同じ（すなはち、 $A_1 < A_2$ ）である。

【0067】3) 音楽片  $A_1, A_2$  の規格化された時間における開始時間  $T_{A_1}, T_{A_2}$  とし、検索キーの音楽片  $K_1, \dots, K_n$  の規格化された時間における開始時間  $T_{K_1}, \dots, T_{K_n}$  としたとき、開始時間の差額が一致、すなはち、 $T_{A_1} - T_{K_1} = T_{A_2} - T_{K_2}$  である。

【0068】検索結果の出力では、検索キーと、検索結果の候補となったデータベース中の楽曲の一部もしくは全体との間の類似度を求めた後、類似度の高い順番に、楽曲の一部もしくは全体を提示する。データベース中の情報から楽曲の題名、楽曲家名を表示することもできるし、検索キーと類似する楽曲の一覧もしくは全体を表示することもできる。

【0069】本実施の形態においては、データベースに流入する楽曲として MIDI データを用いることしかできる。MIDI データは標準化された形式で、非常に多くの楽曲が MIDI データとして存在しており、データベースの構築のために楽曲の演奏を行なうたりする必要がなく、データベース構築が容易である。また、MIDI データは必ず基準として音符情報が記述されているため、テンポが統一化。音高情報の抽出が正確に行なえるため、音楽検索用のデータベースに利用することに適している。

【0070】また、検索キーとなる楽曲の入力の際に、リズムと受け発信楽曲を補助的に使うことによって、検索キーとなる楽曲のテンポが規格化が正確に行なえるようになる。

【0071】本実施の形態の特徴的な効果として、データベースの検索が容易かつ正確に行なえること、また、検索キーとなる楽曲のテンポが規格化が正確に行なえることが挙げられる。

【0072】（第2の実施の形態）本実施の第2の実施の形態においては、上記の第1の実施の形態でデータベースに入れるする楽曲として MIDI データを用いるものではなく、例えば音楽用 MIDI (コンパクト・ディスク) の

データや、WAVE、AIFF、AIFなどの形式のよう  
にPCM符号化された楽曲、もしくは、他の符号化形式  
のものであれば、それを一度PCM符号化形式に変換し  
たものを用いる。

【0073】一般に楽曲においては、音子に合わせて音  
が演奏される。そして、音は、演奏が開始される瞬間に  
大きく音圧が変化するため、楽曲における音圧の変動を  
測れば、テンポ情報を抽出することが可能である。

【0074】より具体的なテンポ情報を抽出方法として、  
以下のようない方法がある。

【0075】① 100Hzや10Hzといった小さな  
サンプリング周波数で音量レベルのピークを探し、ピー  
クの間隔からテンポを抽出する方法。元のPCMデータ  
のサンプリング周波数が、例えば44100Hzであったとすると、1000サンプルずつ足しあわせたデータ  
を作成し、音量レベルのピークを探す、といった方法が  
考えられる。

【0076】② 上記小さなサンプリング周波数で音量  
レベルの変化率の推移を調べ、変化率が急激に増加する  
部分を探し、その急激に増加する部分の間隔からテンポ  
を抽出する方法。

【0077】③ 上記小さなサンプリング周波数のデータ  
に対し、例えば、10秒といった長い区間で周波数分  
析を行い、ピークを探す方法。例えば、2日のところ  
にピークがあれば、この楽曲の1拍の長さは0.5秒で  
あるということになる。

【0078】④ リズムを刻むドラムズなどの楽器の音  
を選択的に抽出し、これらの音の音圧の変動を分析する  
ことにより、テンポ情報を抽出する方法。

【0079】入力された楽曲から音高情報をなどの特徴量  
を生成する方法としては、上記の第1の実施の形態でハ  
ミングや楽器の演奏により入力された楽曲から音高情報  
などを抽出する方法に準じる。ただし、ハミングや楽  
器の楽器の場合と違い、音楽用CDなどのデータでは、  
楽曲の音符が特に分離されていないため、楽曲対象と  
なる音高（音楽に付随する音の高さ）のみを抽出する  
必要がある。そのためには、MFDTや自己相關処理をした  
のちに、音圧の最も高いピークを抽出する方法や、旋鈍  
やリップによる演奏する楽器、もしくは人の歌聲の周波数帯  
のピークのみを抽出するような音域フィルタを使用する  
方法がある。

【0080】本実施の形態の効果は以下のとおりである。  
データベースに入力する楽曲として、MP3よりも  
さらに一般的な音楽符号化方法であるMP4形式により  
符号化された楽曲を利用することにより、データベース  
中に含まれる楽曲の難易度より充実したものにするこ  
とが可能である。また、楽曲の音楽性をよりはんなりせず、  
聴覚的に音楽によって演奏した樂曲などをデータベース中に  
含むことができるようになる。

【0081】【第3の実施の形態】本実施の第3の実施

の形態においては、上記の第1、第2の実施の形態でデータベース中の楽曲、楽曲キーの楽曲を音楽片に分類し、特徴量を生成する際に、音高推移ベクトルを生成するのではなく、対象となる音楽片における音高の分布を求める。この分布をもとに特徴量ベクトルを生成する。

【0082】ある音楽片において、音高が、図8(A)に示すように推移した場合、規格化された時間の一単位を一度数に換算して、図8(B)のような音高分布図を作成し、各音高（もしくは音高の範囲）の度数を並べて音高分布ベクトルを作成することができる。これを、音  
楽情報（特徴量）として検索を行うことができる。

【0083】音楽情報（特徴量）としては、上記のほかに、例えば、以下のものを生成することができる。  
【0084】① 第1の実施の形態で利用した音高推移  
ベクトルA、各要素から、要素の平均値を引いてでき  
たベクトルA'の要素の値の分布。

【0085】② 第1の実施の形態で利用した音高推移  
ベクトルA、各要素から、最初の要素の値を引いてでき  
たベクトルA''の要素の値の分布。

【0086】③ 第1の実施の形態で利用した音高推移  
ベクトルA、各要素から、前の要素の値を引いてでき  
たベクトルA'''の要素の値の分布。

【0087】また、上記各分布において、図8(B)の音高分布図の横軸に相当する値が0に相当する部分の度  
数を0とする場合もある。また、上記各分布ににおいて、  
図8(B)の音高分布図の横軸に相当する値を、定められた範囲内に固定し、範囲を超えたデータを削除する場合、もしくは、その範囲の境界の値に加える場合もある。

【0088】これらの特徴量ベクトルを利用して、第1の  
実施の形態の方法に従い、音楽片検索、音楽片検索を統  
合して類似検索を行う。

【0089】本実施の形態の効果について、第1の実施  
の形態の場合よりも、一般に特徴量ベクトルの次元数を  
減らすことができるため、検索速度を向上させること  
が可能である。また、音楽片の範囲内では、瞬間の情報  
が失われるため、データベース中の音楽片と、検索キ  
ーから生成された音楽片の開始地点に少しの誤差があ  
っても、検索が可能となる。一方、瞬間の情報が失われる  
ため、音高推移ベクトルで見れば似似しないもので  
も、音高分布ベクトルでは類似する可能性がある。した  
かって、まず本実施の形態による方法で検索を行い、そ  
の検索結果の中から、第1の実施の形態による検索を行  
う段階で検索などをを行うことを可能である。

【0090】

【特徴量の抽出】上記のように、音楽片に記載する、複数音  
がハミングや楽器の演奏などにより、自由に楽曲の一  
部または全体を、力をすることにより、複数音が発音した  
樂曲をデータベースの中から複数よく高周波で検索するこ  
とが可能となる。

【009】本発明によって、例えば、カラオケや、ミュージック・オン・デマンドシステム（自動ジャukeボックス）などのシステムにおいて、利用者が選択したい楽曲の名前や作曲者名、演奏家名などを思い出せない場合においても、その楽曲のメロディーさえ覚えていれば目的の楽曲を選択可能となり、このような分野において極めて有効である。

【図面の簡単な説明】

【図1】本発明による音楽の類似検索（音楽情報への変換を行わない場合）の概要を説明する図である。

【図2】本発明による音長の類似検索（音楽情報への変換を行う場合）の概要を説明する図である。

【図3】本発明の原理構成図である。

【図4】テンポ規格化を説明する図である。

【図5】マッチングの単位を説明する図である。

【図6】音楽片への分割を説明する図である。

【図7】音楽情報への変換を説明する図である。

【図8】音高情報への変換（音高分布）を説明する図である。

【図9】従来方法による音楽検索（D Pマッチング使用）の原理説明図である。

\*【図10】本発明に関連する音楽片を使用する音楽検索の原理説明図である。

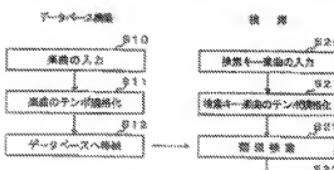
【図11】本発明に関連する音楽情報を使用しない音楽検索の原理説明図である。

【符号の説明】

1	楽曲入力手段
1.1	リズム信号発信手段
1.2	リズム信号受信手段
2	テンポ規格化手段
10 2.1	テンポ抽出手段
3	音楽情報変換手段
3.1	音楽情報変換手段
3.2	音楽片生成手段
4	楽曲蓄積手段
4.1	楽曲蓄積部
4.2	音楽情報蓄積部
4.3	音楽片蓄積部
4.4	インデックス生成手段
5	楽曲検索手段
20 5.1	音楽片検索手段
5.2	音楽片整合手段

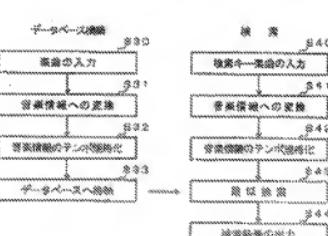
【図1】

本発明による音楽の類似検索（音楽情報への変換を行わない場合）の概要

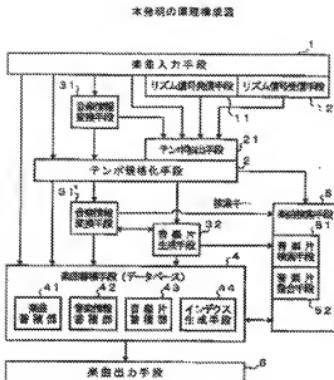


【図2】

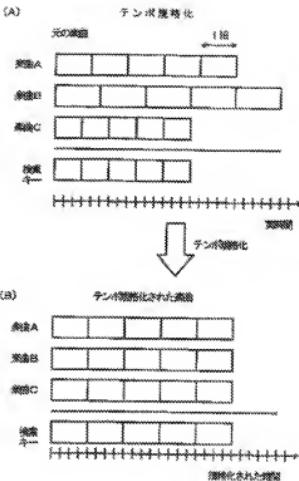
本発明による音楽の類似検索（音楽情報への変換を行う場合）の概要



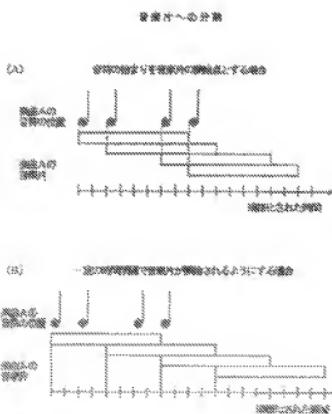
【図3】



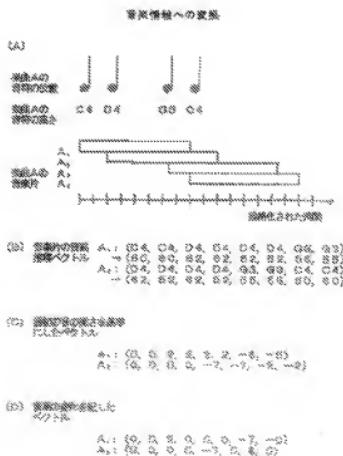
【図4】



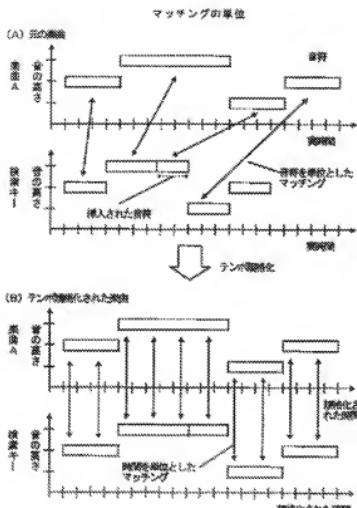
【図5】



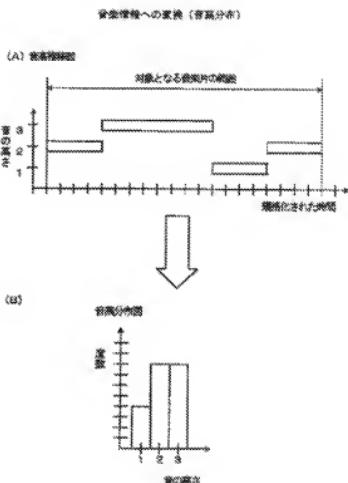
【図7】



200



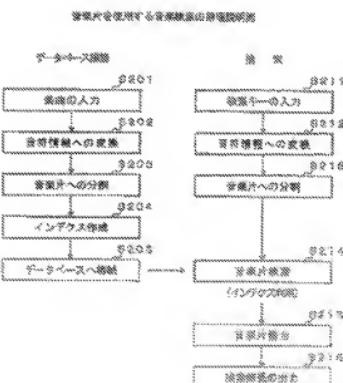
〔卷六〕



200

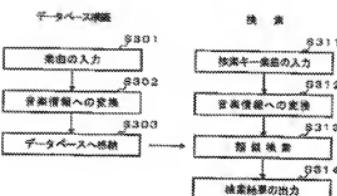


卷之三



【図11】

銀行情報を使用しない音楽検索の実施詳細



## フロントページの続き

(72)発明者 坂田 哲夫

東京都千代田区大手町二丁目3番1号 日  
本電信電話株式会社内

(72)発明者 山塙 雅明

東京都千代田区大手町二丁目3番1号 日  
本電信電話株式会社内

(72)発明者 梅田 昌義

東京都千代田区大手町二丁目3番1号 日  
本電信電話株式会社内

(72)発明者 芝谷 稲一

東京都千代田区大手町二丁目3番1号 日  
本電信電話株式会社内